Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Economic Dynamics & Control 28 (2004) 1461 – 1480
JOURNAL OF Economic Dynamics & Control
www.elsevier.com/locate/econbase

Innovations, improvements, and the optimal adoption of new technologies

Ulrich Doraszelski*

Hoover Institution, Stanford University, Stanford, CA 94305-6010, USA

Abstract

This paper extends the literature on technology adoption by introducing a distinction between technological breakthroughs ('innovations') and the engineering refinements ('improvements') that follow such a breakthrough. Firms do not necessarily wait for a future technological breakthrough, but instead have an incentive to delay the adoption of a new technology until it is sufficiently advanced.

We characterize a firm's adoption decision as the solution to a continuous time, infinite horizon dynamic programming problem, which gives rise to an ordinary differential equation that is highly non-linear and does not have a closed-form solution. We thus employ projection techniques and show how these techniques can be applied to the analysis of optimal stopping problems.

JEL classification: D81; D92; O33

Keywords: Technology adoption; Technological uncertainty; Optimal timing; Investment irreversibility; Waiting option value

1. Introduction

A striking empirical observation regarding the adoption of new technologies is that there is typically a substantial lag between the discovery of a new technology and its adoption. Early work by Schumpeter (1934, 1950) and Mansfield (1968) has attributed the lag in the adoption of a new technology to the uncertainty about the nature of technological change. Models of *uncertain profitability* argue that a firm has an incentive to delay adoption because it can gather information as time passes, and thus perhaps avoid adopting an unprofitable technology (Jensen, 1982; McCardle, 1985; Bhattacharya et al., 1986).

* Corresponding author.
E-mail address: doraszelski@hoover.stanford.edu (U. Doraszelski).

0165-1889/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0165-1889(03)00112-X

Models of *technological uncertainty* extend this line of research by acknowledging that there is, in addition, uncertainty generated by further technological developments. Moreover, if there is rapid technological progress, then there is very little chance that a firm can recover its investment once it has installed a new technology. The adoption decision is thus irreversible. Ongoing technological progress complicates a firm's decision problem considerably because the firm should weigh two types of cost against each other: On the one hand, the cost of making a mistake by adopting too soon and, on the other hand, the opportunity cost of waiting in anticipation of a more efficient future technology. Consequently, the adoption decision depends critically on how fast and by how much the firm expects that technology will advance over time. [1]

Firms' expectations regarding further technological developments coincide with the stochastic process that governs technological change (at least if expectations are rational). Yet, the existing literature has paid little attention to the specification of this stochastic process. This paper is an attempt to correct this deficit. In particular, we contribute to the existing literature by introducing a distinction between technological breakthroughs ('innovations') and the engineering refinements ('improvements') that follow a technological breakthrough. The idea is simple: A technological breakthrough triggers a myriad of small improvements that enhance the efficiency of the basic technology. While each of these improvements is small, they add up to a substantial efficiency gain, thereby making it worthwhile for a firm to wait before adopting the new technology. Hence, firms do not necessarily wait for a future technological breakthrough, but instead have an incentive to delay the adoption of a new technology until it is sufficiently advanced.

However, if the distinction between innovations and improvements is merely one of size, then a model with both innovations and improvements is equivalent to a model with innovations alone. This is no longer the case once we allow the occurrence of the next improvement to depend on the time elapsed since the last innovation. In this way, we capture the idea that "there may be situations were large-scale improvements are confidently expected *after* the introduction of some major innovation" (Rosenberg, 1976, p. 524, his italics). While an innovation spurs a series of improvements, eventually these improvements become less likely. Hence, as time passes, it may become profitable for the firm to adopt a technology that has been available for some time although it was not profitable to do so before. There may then be a non-negligible time lag between the occurrence of an innovation or improvement and its adoption.

Most closely related to our paper are Balcer and Lippman's (1984) and Farzin et al.'s (1998) models of technological uncertainty. In Balcer and Lippman's (1984) model, firms anticipate a sequence of innovations of uncertain profitability and revise their expectations about the occurrence of the next innovation as time passes since the last innovation. Balcer and Lippman (1984) show that it is optimal to adopt the currently available technology if the technology a firm has in place lags behind by more than a certain amount. Moreover, this threshold is non-increasing in the time elapsed since the last innovation if the random time between innovations has a non-increasing hazard rate. Hence, as time passes without technological advances, a firm may be willing to

[1] For a survey of the literature on technology adoption see Reinganum (1989) and Hoppe (2002).

purchase an existing technology even though it was not willing to do so in the past, provided it expects the hazard rate to be non-increasing. However, Balcer and Lippman (1984) fail to justify why the firm should hold this kind of expectations in the first place.

Farzin et al. (1998) resort to a simpler specification of the stochastic process underlying technological change and assume that the interarrival time of innovations is exponentially distributed. Due to the memorylessness property of the exponential distribution, firms' expectations about the occurrence of the next innovation are time invariant, so that the adoption threshold does not vary with the time elapsed since the last innovation. Hence, a new technology is adopted as soon as an innovation drives technological efficiency over a certain threshold while firms leapfrog innovations if technological efficiency remains below this threshold, and there is no room for a time lag between the occurrence of an innovation and its adoption. [2]

Our distinction between innovations and improvements substantially enriches the stochastic process that characterizes technological progress. In particular, we extend Farzin et al.'s (1998) model by building on the idea that the occurrence of the next improvement depends on the time elapsed since the previous innovation. There may then be a non-negligible time lag between the occurrence of an innovation or improvement and its adoption. Unlike Balcer and Lippman (1984), the distinction between innovations and improvements allows us to justify why firms should at times expect a higher speed or extent of technological change.

This distinction is consistent with observations made by de Solla Price (1984) in a comprehensive historical study. He argues that technological progress takes two fundamentally different forms, namely radical shifts at unpredictable times in unpredictable directions towards new technologies and day-to-day improvements of existing technologies. One could also interpret the distinction between innovations and improvements more narrowly. Consider, for example, personal computer systems. There an innovation occurs with the introduction of a new microprocessor design (e.g., Intel 8086, Pentium IV). After being introduced, a basic chip design usually undergoes a series of improvements, most notably in terms of clock speed (Nordhaus, 2001). The point is that both interpretations make it clear that there are important dependencies in the stochastic process that governs technological change. The modeling framework and solution methods developed in this paper accommodate these dependencies.

We describe the firm's adoption decision as the solution to a continuous time, infinite horizon dynamic programming problem. Analyzing the firm's decision problem requires solving an ordinary differential equation (ODE) that characterizes the critical value of technological efficiency. Since this ODE is highly non-linear and does not have a closed-form solution, we employ projection techniques to approximate the unknown function by a high-order polynomial. The basic idea is to choose the coefficients of the polynomial so that the resulting approximation is as close as possible to the true solution of the ODE. To our knowledge, this is the first time that projection methods are applied to optimal stopping problems.

[2] Restricting the stochastic process of technological change even further, Weiss (1994) and Grenadier and Weiss (1997) deal with an installed technology, a currently available innovation, and a future technology.

The remainder of this paper is organized as follows. Section 2 presents the model. Sections 3 and 4 detail the model specification and address a number of computational issues. Section 5 discusses our results. Section 6 concludes with a research agenda.

2. Model

We analyze a dynamic model with infinite time horizon. At time $t=0$ the firm produces with technological efficiency θ_0. The technological efficiency at time t is denoted by θ.

We model a firm's decision to adopt a new technology as a one-time irreversible investment decision. Hence, at time t, the firm can either decide to adopt the current technology or wait. If it adopts it pays a fixed cost $I>0$ and is stuck forever with the adopted technology. Clearly, a more realistic assumption is that the firm can adopt a new technology whenever it is willing to pay for it. We have dealt elsewhere with the case of multiple technology switches (Doraszelski, 2001). There we show that the firm's decision problem is of the same basic form at each technology switch. To simplify the exposition, we therefore restrict attention to the case of a single technology switch in what follows.

Equipped with a technology of efficiency θ the firm makes instantaneous profits $\pi(\theta)$, where π is increasing in its argument. The firm's payoff upon adopting technology θ is thus

$$\int_0^\infty \pi(\theta)\mathrm{e}^{-rs}\,\mathrm{d}s - I = \frac{\pi(\theta)}{r} - I,$$

where $r>0$ is the interest rate. Note that we follow Farzin et al. (1998) in assuming that the fixed cost of installing a new technology I is independent of technological efficiency θ, which is often not unrealistic.[3] On the other hand, extending the model to accommodate a non-constant fixed cost is straightforward.

Innovations as well as improvements enhance the state-of-the-art technology, i.e., we assume that technological change is cumulative. Consider a short interval of time $\mathrm{d}t$. During $\mathrm{d}t$ one of three things can happen: (i) an innovation; (ii) an improvement; (iii) neither of the two. From t to $t+\mathrm{d}t$ technological efficiency evolves according to

$$\mathrm{d}\theta = \begin{cases} U & \text{with probability } \lambda\,\mathrm{d}t, \\ V & \text{with probability } \mu(\tau)\,\mathrm{d}t, \\ 0 & \text{else,} \end{cases}$$

where $U \sim F(u)$ denotes an innovation and $V \sim G(v;\tau)$ an improvement. τ measures the time elapsed since the last innovation and governs current improvements. To prevent technological regress, assume $U>0$ and $V>0$.

[3] For example, in 1982 a state-of-the-art personal computer system was based on an Intel 8086 microprocessor and on a Pentium IV in 2001. The capital cost differs by a factor of 0.35 between these two technologies, whereas the computing power differs by a factor of 11423 (Nordhaus, 2001, Appendix, Table 2). In other words, the capital cost of a personal computer system has not changed much over the past 20 years: At the very least the changes in capital cost pale in comparison to the changes in computing power.

We capture the idea that the speed and/or extent of subsequent improvements depends on the time elapsed since the last innovation by allowing the arrival rate of improvements $\mu(\tau)$ and/or the distribution of improvements $G(v;\tau)$ to depend on τ. Below we explore three scenarios, namely a time-invariant arrival rate, diminishing returns, and learning. In case of a constant arrival rate, τ does not enter the arrival rate of improvements at all. If it enters, however, it seems reasonable to assume that the arrival rate eventually declines due to diminishing returns. In other words, the obvious improvements are realized shortly after the basic technology becomes available. Alternatively, a hump-shaped arrival rate due to learning seems plausible. This reflects the idea that a new technology needs to be understood before it can be improved. Similarly, we distinguish between a time-invariant distribution of improvements, diminishing returns, and learning. In case of diminishing returns, we assume that $G(v;\tau)$ is more "favorable" than $G(v;\tau+\mathrm{d}t)$.[4] Here the idea is that the most promising improvements are realized first. Alternatively, learning might be taken into account.

Let $V(\theta_0,\theta,\tau)$ denote the expected discounted profits of a firm using technology θ_0 when the currently available technology is θ and the last innovation was realized at time $t-\tau$. Since θ_0 does not change throughout the analysis, we drop θ_0 as an argument of V (and subsequently as an argument of the critical value of technological efficiency θ^*). The Bellman equation is

$$\begin{aligned} V(\theta,\tau) = \pi(\theta_0)\,\mathrm{d}t + \frac{1}{1+r\,\mathrm{d}t}\Bigg\{ &\lambda\,\mathrm{d}t \int_0^\infty \max\left(\frac{\pi(\theta+u)}{r} - I, V(\theta+u,0)\right)\mathrm{d}F(u) \\ &+ \mu(\tau+\mathrm{d}t)\,\mathrm{d}t \int_0^\infty \max\left(\frac{\pi(\theta+v)}{r} - I, V(\theta+v,\tau+\mathrm{d}t)\right)\mathrm{d}G(v;\tau+\mathrm{d}t) \\ &+ (1-\lambda\,\mathrm{d}t - \mu(\tau+\mathrm{d}t)\,\mathrm{d}t)V(\theta,\tau+\mathrm{d}t)\Bigg\}. \end{aligned} \tag{1}$$

The Bellman equation adds the current return and the expected future stream of returns, appropriately discounted, under the presumption that future decisions are made optimally. Hence, the term $\pi(\theta_0)\,\mathrm{d}t$ captures the firm's profits from t to $t+dt$ if it decides not to adopt the current technology. The first term in parentheses represents the firm's choice between adopting and waiting upon the arrival of an innovation. Note that τ, the time elapsed since the last innovation, is reset to zero once an innovation occurs. The second term represents the firm's choice between adopting and waiting upon the arrival of an improvement. The third term describes the evolution of V when neither an innovation nor an improvement occurs. Note that we implicitly assume that the firm is able to distinguish between innovations and improvements. While this distinction is clear in many situations, it may be less so in others, at least without the benefit of hindsight.[5]

[4] Read: $G(v;\tau)$ first-order stochastically dominates $G(v;\tau+\mathrm{d}t)$.

[5] Consider again personal computer systems. There an innovation occurs with the introduction of a new microprocessor design which subsequently goes through a series of improvements in terms of clock speed, and we expect the firm to have little trouble distinguishing an innovation from an improvement. Whether it is as easy to tell apart radical shifts towards new technologies from day-to-day improvements of existing technologies in the sense of de Solla Price (1984) is less obvious.

Rearranging the Bellman equation (1), adding and subtracting $V(\theta,\tau+\mathrm{d}t)$ to the left-hand side, dividing by $\mathrm{d}t$, and taking the limit as $\mathrm{d}t\to 0$ yields

$$-\frac{\partial V(\theta,\tau)}{\partial\tau}+(r+\lambda+\mu(\tau))V(\theta,\tau)$$
$$=\pi(\theta_0)+\lambda\int_0^\infty \max\left(\frac{\pi(\theta+u)}{r}-I, V(\theta+u,0)\right)\mathrm{d}F(u)$$
$$+\mu(\tau)\int_0^\infty \max\left(\frac{\pi(\theta+v)}{r}-I, V(\theta+v,\tau)\right)\mathrm{d}G(v;\tau). \tag{2}$$

Following Farzin et al. (1998), we assume that the solution to the functional equation (2) can be characterized by a trigger level above which it is optimal for the firm to adopt the current technology.

Assumption 1. For any $\tau\geqslant 0$ there exists a unique $\theta^*(\tau)$ such that the firm finds it optimal to adopt the current technology if and only if $\theta\geqslant\theta^*(\tau)$.

This assumption is intuitively appealing since it rules out that a firm is willing to adopt a certain technology but is unwilling to adopt a better technology. To better understand how the adoption threshold reflects the distinction between innovations and improvements, suppose that there has just been an advance in technology. If technology has advanced due to an innovation, the critical value of technological efficiency is $\theta^*(0)$, but if technology has advanced because of an improvement, the critical value is $\theta^*(\tau)$, where $\tau>0$ is the time elapsed since the last innovation. As we show below, in general we have $\theta^*(0)\neq\theta^*(\tau)$ whenever $\tau>0$. Hence, the adoption threshold after an innovation has occurred is different from the adoption threshold after an improvement has occurred.

Suppose the current technology has efficiency $\theta^*(\tau)$.[6] Then the firm will adopt immediately after any $U>0$ or $V>0$. Hence, the Bellman equation (2) becomes

$$-\frac{\partial V(\theta^*(\tau),\tau)}{\partial\tau}+(r+\lambda+\mu(\tau))V(\theta^*(\tau),\tau)$$
$$=\pi(\theta_0)+\lambda\int_0^\infty \left(\frac{\pi(\theta^*(\tau)+u)}{r}-I\right)\mathrm{d}F(u)$$
$$+\mu(\tau)\int_0^\infty \left(\frac{\pi(\theta^*(\tau)+v)}{r}-I\right)\mathrm{d}G(v;\tau).$$

Moreover, for $\theta=\theta^*(\tau)$ the firm must be indifferent between adopting and waiting. Hence,

$$V(\theta^*(\tau),\tau)=\frac{\pi(\theta^*(\tau))}{r}-I,$$

[6] Strictly speaking $\theta^*(\tau)-\varepsilon$, where ε is small but positive. Then take limits as $\varepsilon\to 0$.

which implies

$$\frac{\partial V(\theta^*(\tau),\tau)}{\partial \tau} = \frac{1}{r}\pi'(\theta^*(\tau))\theta^{*\prime}(\tau).$$

Plugging in yields the basic differential equation

$$\begin{aligned}
-\frac{1}{r}\pi'(\theta^*(\tau))\theta^{*\prime}(\tau) + (r+\lambda+\mu(\tau))\left(\frac{\pi(\theta^*(\tau))}{r} - I\right) \\
= \pi(\theta_0) + \lambda\int_0^\infty \left(\frac{\pi(\theta^*(\tau)+u)}{r} - I\right)\mathrm{d}F(u) \\
+ \mu(\tau)\int_0^\infty \left(\frac{\pi(\theta^*(\tau)+v)}{r} - I\right)\mathrm{d}G(v;\tau). \qquad (3)
\end{aligned}$$

To obtain a terminal condition for Eq. (3), we assume that the arrival rate and the distribution of improvements eventually settle down.

Assumption 2. The arrival rate and the distribution of improvements approach limits as the time elapsed since the last innovation grows large, $\lim_{\tau\to\infty}\mu(\tau) = \mu$ and $\lim_{\tau\to\infty} G(v;\tau) = G(v)$.

Consequently, $\theta^*(\tau)$ eventually settles down. The following proposition characterizes the steady state.

Proposition 1. *The trigger level approaches a limit as the time elapsed since the last innovation grows large*, $\lim_{\tau\to\infty}\theta^*(\tau) = \theta^*$, *where* θ^* *solves*

$$\begin{aligned}
(r+\lambda+\mu)\left(\frac{\pi(\theta^*)}{r} - I\right) = \pi(\theta_0) + \lambda\int_0^\infty \left(\frac{\pi(\theta^*+u)}{r} - I\right)\mathrm{d}F(u) \\
+ \mu\int_0^\infty \left(\frac{\pi(\theta^*+v)}{r} - I\right)\mathrm{d}G(v). \qquad (4)
\end{aligned}$$

Proof. Take limits on both sides of Eq. (3) and use the fact that $\lim_{\tau\to\infty}\mu(\tau)=\mu$ and $\lim_{\tau\to\infty} G(v;\tau) = G(v)$ implies $\lim_{\tau\to\infty}\theta^*(\tau)/\partial\tau = 0$. □

The final proposition shows that a model with both innovations and improvements is equivalent to a model with innovations alone provided the distribution of improvements and the arrival rate are time invariant.

Proposition 2. *If* $G(v;\tau) = G(v)$ *and* $\mu(\tau) = \mu$ *for all* $\tau \geqslant 0$, *then*, *without loss of generality*, *we can set* $\mu = 0$.

Proof. Redefine λ as $\lambda+\mu$ and $F(u)$ as $[\lambda/(\lambda+\mu)]F(u)+[\mu/(\lambda+\mu)]G(u)$; set $\mu=0$. □

Hence, if the distinction between innovations and improvements is merely one of size, then a model with both innovations and improvements is equivalent to a model with innovations alone. Since the basic differential equation (3) reduces to

$$(r+\lambda)\left(\frac{\pi(\theta^*)}{r}-I\right)=\pi(\theta_0)+\lambda\int_0^\infty\left(\frac{\pi(\theta^*+u)}{r}-I\right)\mathrm{d}F(u), \tag{5}$$

τ is no longer a state variable in the Bellman equation, and the firm finds it optimal to adopt the current technology if and only if $\theta \geqslant \theta^*$. Such a model predicts that a new technology is adopted as soon as an innovation or improvement drives technological efficiency over a certain threshold.

3. Specification

Before we are able to numerically solve the model, we have to specify the stochastic process of technological change. To capture the idea that technology progresses in finite steps, we assume that the support of the distribution of innovations is the compact interval $[\underline{u}, \bar{u}]$, where $\underline{u} \geqslant 0$. An innovation is a random variable U, where

$$U=\tilde{U}(\bar{u}-\underline{u})+\underline{u}$$

and $\tilde{U}$ follows a beta distribution with u_1 and u_2 degrees of freedom. The beta distributions are a fairly flexible family of distributions and include the uniform distribution used by Farzin et al. (1998) as a special case ($u_1 = u_2 = 1$).

We assume that the support of the distribution of improvements is the compact interval $[\underline{v}, \bar{v}]$, where $\underline{v} \geqslant 0$. An improvement is a random variable V, where

$$V=\tilde{V}(\bar{v}-\underline{v})+\underline{v}$$

and $\tilde{V}$ follows a beta distribution with $v_1(\tau)$ and $v_2(\tau)$ degrees of freedom. If $v_1(\tau) < v_2(\tau)$, the distribution of V is skewed to the left and we favor small over large improvements, whereas the distribution of V is skewed to the right and we emphasize improvements with large efficiency gains over those with small ones if $v_1(\tau) > v_2(\tau)$.

We distinguish three cases, namely a time-invariant distribution of improvements, diminishing returns, and learning. A time-invariant distribution is obtained by setting $v_1(\tau) = v_1$ and $v_2(\tau) = v_2$, where $v_1 < v_2$. Diminishing returns in the distribution of improvements are modeled by letting $v_1(\tau)$ decrease with τ while $v_2(\tau)$ increases. This implies that the expected size of improvements is declining over time. In particular, we choose

$$v_1(\tau)=\begin{cases}\dfrac{(v_2-v_1)\cos(\pi\frac{\tau}{\bar{\tau}})+v_1+v_2}{2} & \text{if } \tau \leqslant \bar{\tau},\\[2ex] v_1 & \text{if } \tau > \bar{\tau},\end{cases}$$

$$v_2(\tau)=\begin{cases}\dfrac{(v_2-v_1)\cos(\pi(\frac{\tau}{\bar{\tau}}+1))+v_1+v_2}{2} & \text{if } \tau \leqslant \bar{\tau},\\[2ex] v_2 & \text{if } \tau > \bar{\tau}.\end{cases}$$

Table 1
Parameter values

Parameter	$\underline{u}$	$\bar{u}$	u_1	u_2	λ	$\underline{v}$	$\bar{v}$	v_1	v_2	$\underline{\mu}$	$\bar{\mu}$	$\bar{\tau}$
Value	0	0.2	4	2	0.1	0	0.2	2	4	0.5	2	10

The impact of learning on the distribution of improvements is captured using humped-shaped time paths. We set

$$v_1(\tau) = \begin{cases} \dfrac{(v_1 - v_2)\cos(2\pi\frac{\tau}{\bar{\tau}}) + v_1 + v_2}{2} & \text{if } \tau \leqslant \bar{\tau}, \\ v_1 & \text{if } \tau > \bar{\tau}, \end{cases}$$

$$v_2(\tau) = \begin{cases} \dfrac{(v_1 - v_2)\cos(2\pi(\frac{\tau}{\bar{\tau}} + \frac{1}{2})) + v_1 + v_2}{2} & \text{if } \tau \leqslant \bar{\tau}, \\ v_2 & \text{if } \tau > \bar{\tau}. \end{cases}$$

In contrast to diminishing returns, this implies that the expected size of improvements is first increasing, then decreasing.

Turning to the arrival rate of improvements, we again distinguish three cases. A time-invariant arrival rate is obtained by setting

$$\mu(\tau) = \frac{\underline{\mu} + \bar{\mu}}{2},$$

where $\underline{\mu} < \bar{\mu}$. Diminishing returns are modeled by letting $\mu(\tau)$ decrease with τ,

$$\mu(\tau) = \begin{cases} \dfrac{(\bar{\mu} - \underline{\mu})\cos(\pi\frac{\tau}{\bar{\tau}}) + \underline{\mu} + \bar{\mu}}{2} & \text{if } \tau \leqslant \bar{\tau}, \\ \underline{\mu} & \text{if } \tau > \bar{\tau}, \end{cases}$$

and learning is modeled using a humped-shaped time path

$$\mu(\tau) = \begin{cases} \dfrac{(\bar{\mu} - \underline{\mu})\cos(2\pi(\frac{\tau}{\bar{\tau}} + \frac{1}{2})) + \underline{\mu} + \bar{\mu}}{2} & \text{if } \tau \leqslant \bar{\tau}, \\ \underline{\mu} & \text{if } \tau > \bar{\tau}. \end{cases}$$

The parameters governing the stochastic process of technological change are given in Table 1. The mean time between innovations is $1/\lambda$ or 10 years. On average, an improvement happens every 6 months if $\mu(\tau) = \bar{\mu}$ and every 2 years if $\mu(\tau) = \underline{\mu}$. Innovations and improvements share the same support ($\underline{u} = \underline{v}$ and $\bar{u} = \bar{v}$), but we choose $u_1 > u_2$ and $v_1 < v_2$. Since this skews the distribution of innovations to the right, innovations with large efficiency gains are more likely than those with small ones. In contrast, the distribution of improvements is skewed to the left to favor small over

Table 2
Parameter values

Parameter	α	p	w	I	r	θ_0
Value	0.5	200	50	1600	0.1	1

large improvements.[7] Finally, the effects of diminishing returns and learning last for at most 10 years after an innovation has occurred.

Next we specify the firm's instantaneous profit function. Following Farzin et al. (1998), we consider a price-taking firm that produces a homogeneous good according to the production function

$$y = \theta x^{\alpha}, \quad 0 < \alpha < 1,$$

where θ measures the efficiency of the technology in use. Let p denote the constant price of output and w the constant price of the input. Then the firm's instantaneous profit function is

$$\pi(\theta) = \max_x py - wx = (1-\alpha)\left(\frac{\alpha}{w}\right)^{\alpha/(1-\alpha)} (p\theta)^{1/(1-\alpha)}.$$

The remaining parameters (taken from Farzin et al., 1998) are presented in Table 2. Compared to the net present value of remaining with the technology already in place, $\pi(\theta_0)/r = 2000$, the fixed cost $I = 1600$ of adopting a new technology is substantial.

4. Computation

We numerically solve the model in two steps. First, we solve the non-linear equation (4) to obtain θ^* using the fzero-routine of Matlab 5.2, a bisection algorithm. Second, we employ a projection method to solve the differential equation (3) subject to the terminal condition $\lim_{\tau\to\infty} \theta^*(\tau) = \theta^*$. In particular, we approximate $\theta^*(\tau)$ by a finite sum of Chebyshev polynomials, $\sum_{k=0}^{K} \psi_k \phi_k(\tau)$, where $\phi_k(\tau)$ is a kth-order Chebyshev polynomial in τ. The $K+1$ unknown coefficients $(\psi_0, \ldots, \psi_K)$ are chosen to solve the differential equation at K so-called collocation points and to satisfy the terminal condition $\lim_{\tau\to\infty} \theta^*(\tau) = \theta^*$ (Judd, 1992, 1998, Chapter 11). We solve for $(\psi_0, \ldots, \psi_K)$ using the gradient-based non-linear equation solver c05nbf of the NAG toolbox. Below we discuss these two steps in greater detail.

Both steps employ numerical integration to evaluate the integral in Eqs. (4) and (3), respectively. We numerically integrate using Gauss–Legendre quadrature rather than Matlab's build-in Newton–Cotes routines because, for a given number of function evaluations, the former yields a higher precision (Judd, 1998, Chapter 7). The required

[7] This is in line with technological change in personal computer systems, where new microprocessor designs (i.e., innovations) tend to be associated with large efficiency gains relative to the subsequent improvements in clock speed. The results in Section 5 remain qualitatively the same if we choose $u_1 \leqslant u_2$ instead of $u_1 > u_2$.

Table 3
Maximum relative error of numerical integration

	8 nodes	16 nodes	32 nodes
$\alpha = 0.1$	3.9044×10^{-4}	3.0785×10^{-14}	3.1126×10^{-14}
$\alpha = 0.5$	3.3849×10^{-4}	2.7822×10^{-14}	2.7705×10^{-14}
$\alpha = 0.9$	2.4277×10^{-4}	4.6024×10^{-14}	4.6195×10^{-14}

nodes and weights are supplied by the d01bbf-routine of the NAG toolbox. Note that the numerical integration is the only part of our computational strategy that depends on the specific distributions of innovations and improvements. Other distributions are readily accommodated by changing the quadrature method. [8]

To assess the accuracy of the Gauss–Legendre quadrature, we compute $\int_0^\infty \pi(\theta_0+u)\,\mathrm{d}F(u)$ for $\alpha \in \{0.1, 0.5, 0.9\}$ and $u_1, u_2 \in \{1, 2, \ldots, 10\}$ with 8, 16, and 32 nodes. Maple 5.1 provides an analytic expression for this integral (in terms of beta, gamma, and generalized hypergeometric functions). The maximum error over $u_1, u_2 \in \{1, 2, \ldots, 10\}$ relative to the exact solution is given in Table 3. Since the precision is high for 16 and more nodes, we use 16 nodes henceforth. The high precision is probably due to the polynomial structure of both π and f, where f denotes the density corresponding to F.

Given our functional forms for $\mu(\tau)$ and $G(v;\tau)$, $\mu = \lim_{\tau\to\infty} \mu(\tau) = \mu(\bar{\tau})$ and $G(v) = \lim_{\tau\to\infty} G(v;\tau) = G(v;\bar{\tau})$. Since the effects of diminishing returns and/or learning have been fully realized after $\bar{\tau}$ years, $\theta^*(\tau)$ not only approaches the steady state θ^* as τ grows large but $\theta^*(\tau)$ equals θ^* for $\tau \geqslant \bar{\tau}$. This allows us to rewrite Eq. (4) as $\Delta = 0$, where

$$\Delta = \frac{1}{\pi(\theta_0)} \left\{ (r + \lambda + \mu(\bar{\tau})) \left(\frac{\pi(\theta^*)}{r} - I \right) \right. - \pi(\theta_0) - \lambda \int_0^\infty \left(\frac{\pi(\theta^* + u)}{r} - I \right) \mathrm{d}F(u) \left. - \mu(\bar{\tau}) \int_0^\infty \left(\frac{\pi(\theta^* + v)}{r} - I \right) \mathrm{d}G(v;\bar{\tau}) \right\}.$$

Note that dividing by $\pi(\theta_0)$ renders the residual unit free. Δ is increasing in θ^* provided that π is convex in its argument. Thus, given our model specification, there exists a unique solution to $\Delta = 0$. Moreover, it must be the case that $\theta^* > \theta_0$ since it cannot be optimal for the firm to adopt a less efficient technology than the one it currently uses. To bracket the solution to $\Delta = 0$, we compare the function value at $\theta_0 \times 10^i$ with the one at $\theta_0 \times 10^{i+1}$ for $i = 0, 1, \ldots$ until we find an interval containing a sign change. Given these initial values, the above equation is solved using Matlab's fzero with a precision of 10^{-12}. This bisection algorithm converges quickly for all parameter values.

[8] For example, a lognormal distribution calls for Gauss–Laguerre quadrature (Judd, 1998, Chapter 7).

Eq. (3) can be rewritten as $\Delta(\tau) = 0$, where

$$\Delta(\tau) = \frac{1}{\pi(\theta_0)} \left\{ -\frac{1}{r}\,\pi'(\theta^*(\tau))\theta^{*\prime}(\tau) + (r + \lambda + \mu(\tau)) \left(\frac{\pi(\theta^*(\tau))}{r} - I \right) \right.$$
$$- \pi(\theta_0) - \lambda \int_0^\infty \left(\frac{\pi(\theta^*(\tau) + u)}{r} - I \right) \mathrm{d}F(u)$$
$$\left. - \mu(\tau) \int_0^\infty \left(\frac{\pi(\theta^*(\tau) + v)}{r} - I \right) \mathrm{d}G(v;\tau) \right\}.$$

Note that $\Delta(\tau)$ is again unit free. Since $\theta^*(\tau)$ equals θ^* for $\tau \geqslant \bar{\tau}$, it remains to solve this non-linear non-autonomous differential equation for $0 \leqslant \tau < \bar{\tau}$. To this end we approximate $\theta^*(\tau)$, $0 \leqslant \tau < \bar{\tau}$, by a finite sum of Chebyshev polynomials

$$\hat{\theta}^*(\tau) = \sum_{k=0}^{K} \psi_k \phi_k(\tau),$$

where $\phi_k(\tau)$ is a kth-order Chebyshev polynomial in τ and $(\psi_0, \ldots, \psi_K)$ is a vector of $K + 1$ unknown coefficients. K determines the order of the approximation. Substituting $\hat{\theta}^*(\tau)$ for $\theta^*(\tau)$ and $\hat{\theta}^{*\prime}(\tau)$ for $\theta^{*\prime}(\tau)$ yields a family of non-linear equations in $(\psi_0, \ldots, \psi_K)$. $(\psi_0, \ldots, \psi_K)$ are chosen to satisfy $\Delta(\tau) = 0$ at K collocation points $0 < \tau_1 < \cdots < \tau_K < \bar{\tau}$ along with the terminal condition $\theta^*(\bar{\tau}) = \theta^*$. The collocation points in turn are the zeros of a Kth-order Chebyshev polynomial adapted to the interval $[0, \bar{\tau}]$ and are given by

$$\tau_k = \left(1 - \cos\left(\frac{2k-1}{2K}\pi\right)\right)\left(\frac{\bar{\tau}}{2}\right), \quad k = 1, \ldots, K.$$

We solve the resulting $K+1$ non-linear equations for the $K+1$ unknown coefficients using the c05nbf-routine of the NAG toolbox with a precision of 10^{-12}. The equation solver converges quickly for all parameter values.

Fig. 1 plots the residuals $\Delta(\tau)$ for $K \in \{6, 12, 18\}$ in case of diminishing returns for the arrival rate and the distribution of improvements (panels 1–3) and in case of learning for the arrival rate and the distribution of improvements (panels 4–6). The latter case turns out to be the most troublesome for the Chebyshev collocation and thus provides a lower bound on the degree of accuracy. As the order of the approximation increases, the approximation includes more terms. Since this increases its flexibility, the size of the residuals decreases in K. Moreover, the residuals oscillate around zero but do not quite exhibit the equioscillation property necessary for a best (with respect to the sup-norm) polynomial approximation (Judd, 1998, Chapter 6). Nevertheless, the approximation appears to be reasonable for K=12. We therefore set K=12 in the following.

5. Results

Our model distinguishes between innovations and improvements. To assess the effect of this distinction on a firm's decision to adopt a new technology, we first present

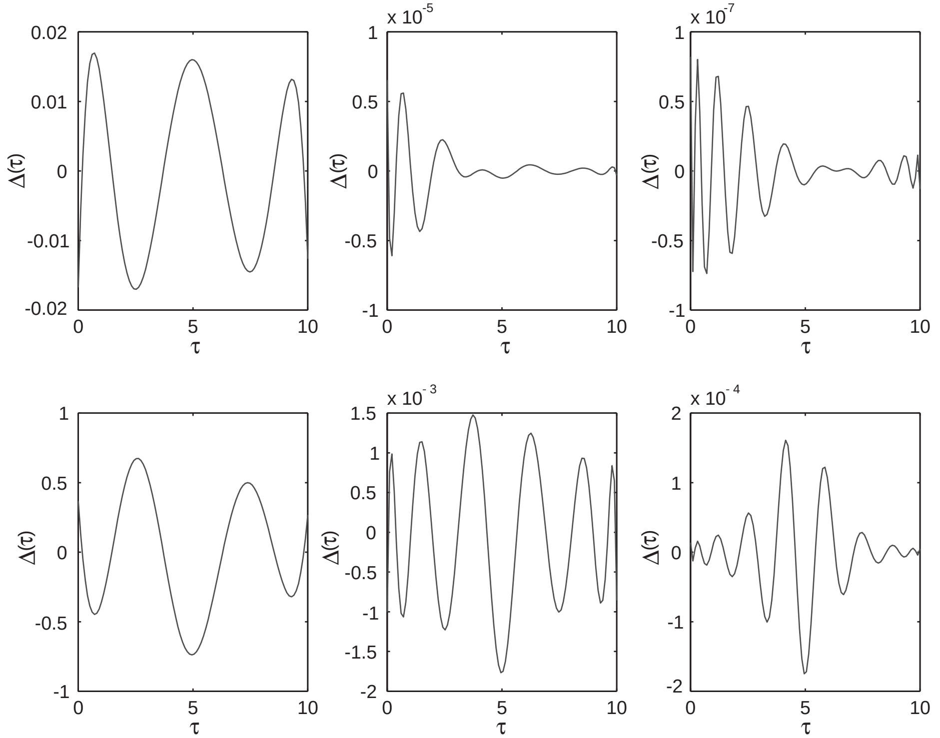

Fig. 1. Residuals $\Delta(\tau)$. Left column: $K=6$, middle column: $K=12$, right column: $K=18$. Top row: diminishing returns for the arrival rate and the distribution of improvements, bottom row: learning for the arrival rate and the distribution of improvements.

some results for the special case without improvements and then turn to the general case.

The stochastic process generating improvements is shut down by setting $\underline{\mu} = \bar{\mu} = 0$. Consequently, the basic differential equation (3) reduces to Eq. (5). As in Farzin et al. (1998), τ is no longer a state variable in the Bellman equation, and the firm finds it optimal to adopt the current technology if and only if $\theta \geqslant \theta^*$. For our parameter values $\theta^*=1.4886$. Fig. 2 presents comparative statics results.

The first panel shows that θ^* increases with λ, since a higher arrival rate renders the immediate availability of a more efficient technology more likely and hence provides the firm with an incentive to delay. In other words, the incentive to postpone adoption increases with the pace of technological progress. It is tempting to conclude that "in conjunction with an increase in the pace (...) of technological innovation we might expect the actual implementations to be not only larger but also less frequent" (Balcer and Lippman, 1984, p. 307).[9] However, while firms take larger steps up the technology

[9] The arrival rate in our model is similar to the discovery potential in Balcer and Lippman's (1984) model.

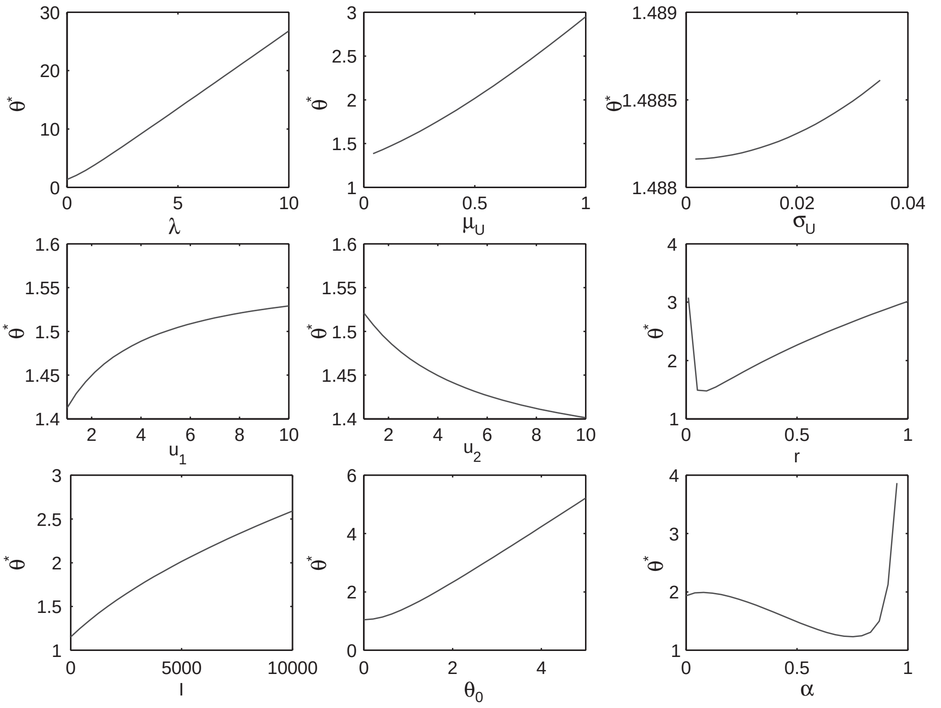

Fig. 2. Comparative statics. Change in θ^* in response to a change in λ, μ_U, σ_U, u_1, u_2, r, I, θ_0, and α. Special case without improvements.

ladder, technology also evolves more rapidly. A simulation shows that the mean time to adoption of 30.67 years for $\lambda = 0.1$ decreases to 16.41 years for $\lambda = 1$ and then increases again to 19.17 years for $\lambda = 10$. Hence, firms may delay or accelerate the adoption of a new technology in response to a change in the arrival rate.

Rather than treating θ^* as a function of $\bar{u}$ alone as in Farzin et al. (1998), we treat θ^* as a function of $\underline{u}$ and $\bar{u}$. This allows us to change the expected value of innovations while holding their standard deviation constant and vice versa. The results are illustrated in panels 2 and 3 of Fig. 2. In line with our intuition, a higher expected value increases the option value of waiting. However, a higher standard deviation increases the option value of waiting as well, albeit by a considerably smaller amount. In other words, the mere possibility of a large efficiency gain is sufficient to induce the firm to delay the adoption of a new technology.

Panels 4 and 5 of Fig. 2 illustrate the dependency of θ^* on u_1 and u_2. Recall that for $u_1 > u_2$, the distribution of innovations is skewed to the right, whereas it is skewed to the left for $u_1 < u_2$, and symmetric for $u_1 = u_2$. Hence, for a given u_2, θ^* increases in u_1, as a higher value of u_1 puts more weight on large efficiency gains and θ^* decreases in u_2 when u_1 is held constant.

Panel 6 of Fig. 2 illustrates that, for small values of r, the trigger value will be lower the higher the discount rate, because a higher discount rate lowers the value of payoffs from more efficient, but also more distant, future technologies, and therefore reduces the option value of waiting. For large values of r, however, the trigger value will be higher the higher the discount rate, since a more efficient technology is required to cover the capital cost associated with its adoption. [10]

Clearly, a higher fixed cost of installing a new technology must be balanced by a greater efficiency gain from the new technology. Panel 7 of Fig. 2 confirms that θ^* is indeed increasing in I. [11]

As can be seen from panel 8 of Fig. 2, the more efficient the technology the firm currently uses, the higher will be the trigger value. This is not surprising because with a relatively efficient technology in place, the gains from adopting a new technology are small compared to the fixed cost the firm incurs. Hence, technology adoption is likely to be slower for firms which are already at the cutting edge of technological efficiency than for those whose current technology lags behind.

The last panel shows θ^* as a function of α. Output elasticity can be interpreted as an indicator of the firm's production efficiency. For relatively large values of α, θ^* goes up with α, implying that the higher a firm's production efficiency, the slower will be the pace of technology adoption. However, for relatively small values of α, the trigger value goes down with the output elasticity. Hence, like firms with high production efficiency, firms with low production efficiency tend to be slow in adopting a new technology, thus suggesting a "low-efficiency technology trap" (Farzin et al., 1998, p. 796). [12]

We now turn to the general case. Fig. 3 presents the time path $\theta^*(\tau)$ along with its value θ^* in steady state. We consider time invariance, diminishing returns, and learning, for both the distribution of improvements and for their arrival rate, yielding a total of nine scenarios.

In line with Proposition 2 the case of a constant distribution of improvements in conjunction with a constant arrival rate is equivalent to the special case without improvements in the sense that the firm finds it optimal to adopt a new technology if and only if $\theta \geqslant \theta^*$, i.e., as soon as an innovation or an improvement drives technological efficiency over the trigger value (panel 1).

This changes once we introduce the idea that the occurrence of the next improvement depends on the time elapsed since the last innovation either through the distribution of improvements or their arrival rate. While the time path of $\theta^*(\tau)$ is monotonically decreasing towards θ^* in case of diminishing returns (panels 2, 4, and 5), $\theta^*(\tau)$ follows a hump-shaped path in the presence of learning (panels 3, 6, 7, 8, and 9). Learning reflects the idea that a new technology needs to be understood before it can be

[10] This is overlooked by Farzin et al. (1998) who claim that the trigger value is monotonic in the discount rate (p. 793). Given their parameter values, θ^* starts to rise for $r \geqslant 0.25$ while they plot θ^* as a function of r for $0 \leqslant r \leqslant 0.2$.

[11] Changing p and w is qualitatively similar to changing I since it is always possible to normalize either the 'scale' of the profit function or the fixed cost to unity by choosing appropriate monetary units.

[12] This result is not an artifact of our model specification and continues to hold for the more general profit function $\pi(\theta) = \alpha_0 \theta^{\alpha_1}$, where $\alpha_0 > 0$ and $\alpha_1 > 0$ is the elasticity of the profit function with respect to technological efficiency. Our results regarding the interest rate and the standard deviation of innovations depend on our model specification: θ^* is increasing in r and decreasing in σ_U for this functional form.

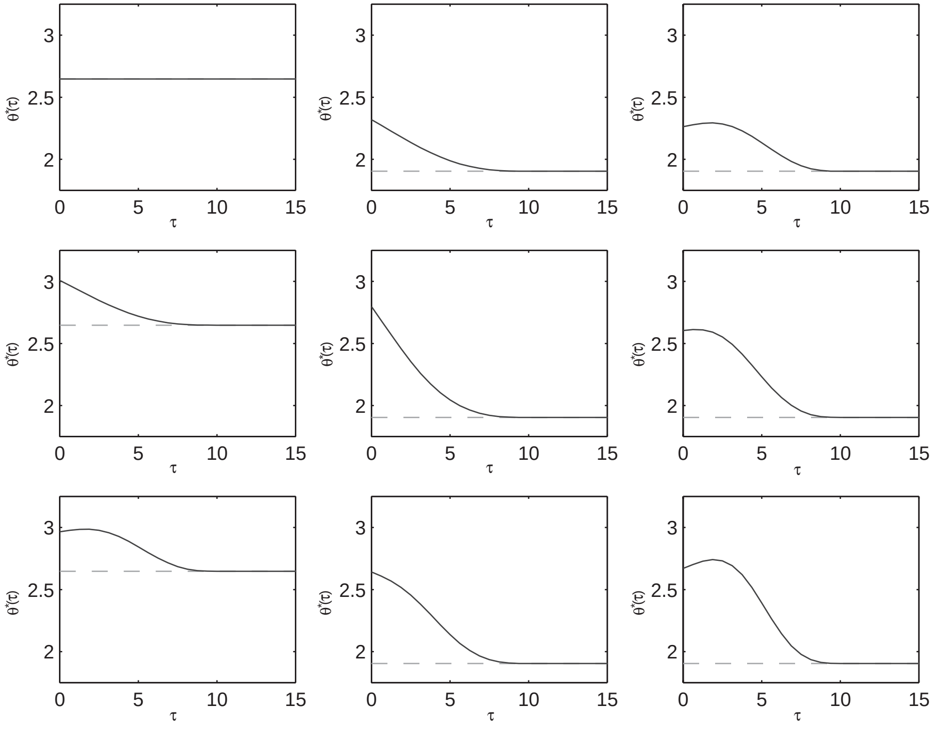

Fig. 3. Time path $\theta^*(\tau)$ (solid line) and steady state θ^* (dashed line). Left column: constant arrival rate, middle column: diminishing returns, right column: learning, top row: constant distribution of improvements, middle row: diminishing returns, bottom row: learning.

improved. Hence, after an innovation, the industry becomes 'ripe' for an improvement as time passes. Consequently, a technology that is not worth adopting today will be even less so tomorrow. Once the obvious improvements have been realized and diminishing returns set in, the likelihood of further advances in technology diminishes with the passage of time, and it may no longer be optimal to defer adoption of the technology.

Since the distribution of improvements and the arrival rate eventually settle down, $\theta^*(\tau)$ eventually decreases towards θ^*. Once $\theta^*(\tau)$ equals θ^*, the firm again finds it optimal to adopt a new technology as soon as an innovation or an improvement drives technological efficiency over the trigger value. In the meantime, however, as time passes it may become profitable for a firm to adopt a technology that has been available for some time although it was not profitable to do so before. This opens the possibility for a time lag between the occurrence of an innovation or improvement and its adoption.[13] Tables 4 and 5 present simulation results on the time to adoption, t^*,

[13] The time lag is finite since θ^* is finite and technology advances in finite time (with probability one). Hence, it is impossible that a firm waits forever.

Table 4
Simulation results

	Without improvements			Constant		
	t^*	τ_{t^*}	lag	t^*	τ_{t^*}	lag
mean	41.8723	0	0	17.4315	9.0398	0
std. dev.	20.4591	0	0	3.9837	10.4772	0
min.	3.3079	0	0	8.4176	0.0000	0
max.	124.7052	0	0	32.1092	63.1209	0

Time to adoption, time elapsed since the last innovation, and lag between the occurrence of an innovation or improvement and its adoption. Left panel: without improvements, right panel: time-invariant distribution of improvements and arrival rate.

Table 5
Simulation results

	Diminishing returns			Learning		
	t^*	τ_{t^*}	lag	t^*	τ_{t^*}	lag
mean	11.0262	5.7386	0.1479	11.7189	7.4841	0.2000
std. dev.	6.9813	9.0184	0.3371	6.3297	8.0523	0.4782
min.	2.3149	0.0000	0.0000	1.3492	0.0000	0.0000
max.	42.1604	66.2675	2.4361	44.8476	77.0071	3.9789

Time to adoption, time elapsed since the last innovation, and lag between the occurrence of an innovation or improvement and its adoption. Left panel: diminishing returns for the distribution of improvements and the arrival rate, right panel: learning for the distribution of improvements and the arrival rate.

the time elapsed since the last innovation τ_{t^*}, and the lag between the occurrence of an innovation or improvement and its adoption.

In the special case without improvements, the mean time to adoption is 41.87 years. Allowing improvements from a time-invariant distribution to arrive at a constant rate increases θ^* from 1.4886 to 2.6474 but decreases the mean time to adoption to 17.43 years. Since the stochastic process is more 'favorable' with both innovations and improvements than with innovations alone,[14] firms are not only more inclined to defer adoption, but technology also evolves more rapidly. As the latter effect dominates the former, firms accelerate the adoption of a new technology. This is an artifact of our parameter values: If improvements are smaller, $\underline{v}=0$, $\bar{v}=0.01$, and arrive less frequently, $\underline{\mu}=0.00625$, $\bar{\mu}=0.025$, then the mean time to adoption increases to 42.13 years. For our parameter values, the distribution of the time to adoption exhibits considerable volatility and is thus able to accommodate a wide variety of empirical patterns. Although firms adopt a new technology on average 9.04 years after an innovation occurred when improvements are added to the model, adoption occurs as soon as an innovation or an improvement drives θ over θ^*.

[14] Read: For any t the distribution of θ with improvements first-order stochastically dominates the distribution of θ without improvements.

Going from Table 4 to Table 5, the mean time to adoption drops further to 11.03 (11.72) years in case of diminishing returns (learning) for the distribution of improvements and the arrival rate. However, adoption of a new technology occurs on average 5.74 (7.48) years after an innovation occurred, and firms adopt a new technology on average 0.15 (0.20) years after an innovation or improvement occurred. Indeed, in 30% (34%) of the simulation runs the firm adopted a new technology not because an innovation or improvement drove technological efficiency over the trigger value but because $\theta^*(\tau)$ dropped below θ in the absence of new discoveries. Conditional on this event, the time lag is 0.49 (0.59) years. In short, letting the occurrence of the next improvement depend on the time elapsed since the last innovation leads to a non-negligible time lag between the occurrence of an innovation or improvement and its adoption.

To gauge how the adoption threshold changes with the parameter values, we again conduct comparative statics experiments. In case of a time-invariant distribution of improvements and arrival rate, in case of diminishing returns, and in case of learning, θ^* behaves just as in the special case without improvements (recall Fig. 2). Clearly, since $\theta^*(\tau) \geqslant \theta^*$, the behavior of θ^* is also indicative of the behavior of $\theta^*(\tau)$. Moreover, the conducted experiments show that $\theta^*(\tau)$ continues to be monotonically decreasing towards θ^* in case of diminishing returns, while it follows a hump-shaped path in the presence of learning irrespective of the parameter values. Overall, our main conclusions remain qualitatively unchanged.

6. Conclusion

We extend the literature on technology adoption by distinguishing between innovations and improvements. The possibility of further improvements gives a firm an incentive to delay the adoption of a new technology until it is sufficiently advanced. The distinction between innovations and improvements therefore allows us to formalize Rosenberg's (1976) insight that "as soon as we accept the perspective of ongoing nature of much technological change, the optimal timing of an innovation (adoption) becomes heavily influenced by expectations concerning the timing and the significance of *future* improvements (...) Thus, a firm may be unwilling to introduce the new technology if it seems highly probable that future technological improvements will shortly be forthcoming" (p. 525, his italics).

We show that a model with both innovations and improvements is equivalent to a model with innovations alone in the sense that a new technology is adopted immediately after an innovation or improvement drives technological efficiency over a certain threshold. But once we allow the distribution of improvements and/or their arrival rate to depend on the time elapsed since the last innovation, the critical value of technological efficiency is no longer constant but instead changes over time. We use projection methods to solve the ODE that characterizes this threshold. The resulting approximation seems to be close to the true solution of the ODE, and gives an accurate description of the solution of the firm's dynamic programming problem. Our results suggest that projection techniques are promising tools for the analysis of optimal stopping problems.

Besides being ubiquitous in financial economics (Dixit and Pindyck, 1994), optimal stopping problems arise from a wide variety of questions, e.g., when to cut a tree, when to replace a durable good, and when to quit a job. To ensure tractability, the existing models are often forced to drastically simplify the stochastic process that governs the state variable. The numerical techniques developed in this paper could be used to allow for hitherto neglected dependencies in this stochastic process, thereby substantially enriching the existing models.

Our computational strategy allows us to provide new insights into a firm's decision to adopt a new technology. In particular, until the effects of diminishing returns and/or learning have been fully realized, there may be a non-negligible time lag between the occurrence of an innovation or improvement and its adoption. Whether the stochastic process of technological change is best described by diminishing returns or learning is an empirical question that needs to be addressed in future research along with the estimation of the values of the parameters governing the stochastic process of technological change.

In our model technological progress is 'manna from heaven'. While this exogeneity assumption seems plausible in the case of technological breakthroughs, it seems less so in the case of engineering refinements. But if engineering refinements are realized while the basic technology is actually in use and if the realized refinements are public knowledge, then there is a positive externality associated with adopting a new technology. This in turn gives each firm an incentive to delay adopting the new technology. On the other hand, if all firms wait, each firm waits in vain, in which case it would have been better to adopt earlier and avoid the opportunity cost of waiting. In other words, firms are involved in a war of attrition (Hendricks and Kovenock, 1989; Hendricks and Porter, 1996), in which each firm has to take its rivals' adoption decisions into account when it makes its own adoption decision.

Strategic interactions also become crucial when technology itself is a source of market power. Specifically, if firms compete in the product market in such a way that each firm's instantaneous profits are declining once one of its competitors adopts a new technology, then each firm has an incentive to itself adopt the new technology as soon as it becomes available in order to forestall being preempted by another firm. While the connection between negative externalities and technology adoption has been studied extensively (see Reinganum (1981a, 1981b) and Fudenberg and Tirole (1985) among others), the existing models along this line generally abstract from uncertain profitability and restrict attention to a single new technology. Hence, these models fail to capture the technological uncertainty that is inherent in the ongoing nature of technological progress.

Acknowledgements

This paper is based on a chapter of my Northwestern University Ph.D. dissertation. An earlier version of this paper was circulated under the title 'Optimal Technology Adoption: The Case of the U.S. Steel Industry'. I would like to thank David Besanko, Michaela Draganska, Dino Gerardi, Heidrun Hoppe, Dipak Jain, Ken Judd, Rosa Matzkin, Rob Porter, Mark Satterthwaite, and two anonymous reviewers for their comments and suggestions.

References

Balcer, Y., Lippman, S., 1984. Technological expectations and adoption of improved technology. Journal of Economic Theory 34, 292–318.

Bhattacharya, S., Chatterjee, K., Samuelson, L., 1986. Sequential research and the adoption of innovations. Oxford Economics Papers 38, 219–243.

de Solla Price, D., 1984. The science/technology relationship, the craft of experimental science, and policy for the improvement of high technology innovation. Research Policy 13, 3–20.

Dixit, A., Pindyck, R., 1994. Investment Under Uncertainty. Princeton University, Princeton, NJ.

Doraszelski, U., 2001. The net present value method versus the option value of waiting: a note on Farzin, Huisman, and Kort (1998). Journal of Economic Dynamics and Control 25, 1109–1115.

Farzin, Y., Huisman, K., Kort, P., 1998. Optimal timing of technology adoption. Journal of Economic Dynamics and Control 22, 779–799.

Fudenberg, D., Tirole, J., 1985. Preemption and rent equalization in the adoption of new technology. Review of Economic Studies 52, 383–401.

Grenadier, S., Weiss, A., 1997. Investment in technological innovations: an option pricing approach. Journal of Financial Economics 44, 397–416.

Hendricks, K., Kovenock, D., 1989. Asymmetric information, information externalities, and efficiency: the case of oil exploration. Rand Journal of Economics 20, 164–182.

Hendricks, K., Porter, R., 1996. The timing and incidence of exploratory drilling on offshore wildcat tracts. American Economic Review 86, 388–407.

Hoppe, H., 2002. The timing of new technology adoption: theoretical models and empirical evidence. Manchester School 70, 56–76.

Jensen, R., 1982. Adoption and diffusion of an innovation of uncertain profitability. Journal of Economic Theory 27, 182–192.

Judd, K., 1992. Projection methods for solving aggregate growth models. Journal of Economic Theory 58, 410–452.

Judd, K., 1998. Numerical Methods in Economics. MIT, Cambridge, MA.

Mansfield, E., 1968. Industrial Research and Technological Innovation. Norton, New York.

McCardle, K., 1985. Information aquisition and the adoption of new technology. Management Science 31, 1372–1389.

Nordhaus, W., 2001. The progress of computing. Discussion Paper No. 1324, Cowles Foundation, New Haven.

Reinganum, J., 1981a. Market structure and the diffusion of new technology. Bell Journal of Economics 12, 618–624.

Reinganum, J., 1981b. On the diffusion of new technology: a game theoretic approach. Review of Economic Studies 48, 395–405.

Reinganum, J., 1989. The timing of innovation: research, development, and diffusion. In: Schmalensee, R., Willig, R. (Eds.), Handbook of Industrial Organization. North-Holland, Amsterdam.

Rosenberg, N., 1976. On technological expectations. Economic Journal 86, 523–535.

Schumpeter, J., 1934. The Theory of Economic Development. Harvard University, Cambridge, MA.

Schumpeter, J., 1950. Capitalism, Socialism, and Democracy, 3rd Ed. Harper & Row, New York.

Weiss, A., 1994. The effects of expectations on technology adoption: some empirical evidence. Journal of Industrial Economics 42, 341–360.